EXHIBIT A

UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

THE NEW YORK TIMES COMPANY, Plaintiff, v. MICROSOFT CORPORATION, OPENAI, INC., OPENAI LP, OPENAI GP, LLC, OPENAI, LLC, OPENAI OPCO LLC, OPENAI GLOBAL LLC, OAI CORPORATION, LLC, and OPENAI HOLDINGS, LLC, Defendants.	Civil Action No. 1:23-cv-11195

THE NEW YORK TIMES COMPANY'S FIRST SET OF REQUESTS FOR PRODUCTION TO OPENAI, INC., OPENAI LP, OPENAI GP, LLC, OPENAI, LLC, OPENAI OPCO LLC, OPENAI GLOBAL LLC, OAI CORPORATION, LLC, AND <u>OPENAI HOLDINGS, LLC</u>

Pursuant to the Federal Rules of Civil Procedure and the Local Rules of the Southern District of New York, (the "Local Civil Rules"), The New York Times Company ("NYT") requests that OpenAI, Inc., OpenAI LP, OpenAI GP, LLC, OpenAI, LLC, OpenAI OPCO LLC, OpenAI Global LLC, OAI Corporation, LLC, and OpenAI Holdings, LLC ("OpenAI Defendants") produce the requested documents at the offices of Susman Godfrey L.L.P., 1301 Avenue of the Americas, 32nd Floor, New York, NY 10019, within thirty days after the date of service of this document, or at such other mutually agreeable time and place.

I.   DEFINITIONS

The following definitions and instructions apply to the requests listed below:

1.   All definitions and rules of construction set forth in Rule 26.3 of the Local Rules of the Southern District of New York are incorporated by reference herein.

2.   "Affiliate" means, with respect to any Person, any other Person that, directly or indirectly through one or more intermediaries, controls, is controlled by, or is under any common control with, such Person.

3.   "Agreement(s)" means any contract, agreement, arrangement, or understanding, formal or informal, oral or written, between or among two or more Persons.

4.   "AI Model(s)" means a component of an information system and/or large language model that implements artificial intelligence ("AI") technology and uses computational, statistical, or machine-learning techniques to produce outputs from a given set of inputs.

5.   "Generative AI Models(s)" means Your AI Models that emulate the structure and characteristics of input data in order to generate derived synthetic content, such as images, videos, audio, text, and other digital content.

6.   "Text Generation AI Model(s)" means Your Generative AI Models, including generative pre-trained transformers or large language models (LLMs), that provide text outputs in response to inputs or prompts, including but not limited to GPT base, GPT-1, GPT-2, GPT-3, GPT 3.5, GPT-4, GPT-4 Turbo and all other versions or other Generative AI Models included in Your Generative AI Products and Services.

7.   "Complaint" means the operative complaint in the Action as of the date these Requests are served.

8.   "Defendants" means Microsoft and OpenAI.

9.   "Electronically Stored Information" or "ESI" refers to information and Documents within the full scope of Federal Rule of Civil Procedure 34 with all Metadata intact—created, manipulated, communicated, stored, and best utilized in digital form, and stored on electronic media.

10.   "Generative AI Product(s) and Service(s)" means any publicly available or commercial product or service offering that includes a Generative AI Model, including but not limited to ChatGPT branded products and services, Copilot branded products and services, ChatGPT Browse with Bing, Bing Chat, APIs, and platforms otherwise providing access to a Generative AI Model.

11.   "Including" means including but not limited to.

12.   "Journalism," for the purposes of these requests, means the activity of writing or creating content for newspapers, magazines, news websites, mobile applications, television, podcasts, or any other publication and/or news outlet, and includes the work of The Times as alleged in the Complaint.

13.   "Metadata" refers to structured information about an electronic file that is embedded in the file, describing the characteristics, origins, usage, and validity of the electronic file.

14.   "OpenAI" or "You" or "Your" means Defendants OpenAI Inc., OpenAI GP, LLC, OpenAI Opco, LLC, OpenAI Global, LLC, OpenAI Holdings LLC, OpenAI, LLC, or OAI Corporation, LLC, their subsidiaries, affiliates, divisions, predecessors in interest, successors, or assignees, and their respective officers, directors, employees, consultants, representatives, and agents.

15.   "Relating to" means discussing, describing, referring to, pertaining to, containing, analyzing, studying, reporting on, commenting on, evidencing, constituting, setting forth, considering, recommending, concerning, or pertaining to, in whole or in part.

16.   "Microsoft" means Defendant Microsoft Corporation, its subsidiaries, affiliates, divisions, predecessors in interest, successors, or assignees, and their respective officers, directors, employees, consultants, representatives, and agents.

17.   "The Times" means Plaintiff The New York Times Company, its subsidiaries, affiliates, divisions, predecessors in interest, successors, or assignees, and their respective officers, directors, employees, consultants, representatives, and agents.

18.   "Times Content" means material published by The Times, including in print, online, audio, or in any other format, whether on Times-owned platforms or available through third-party sources.

19.   "Training Dataset(s)" means the data and content used to train or fine tune the Generative AI Models.

II.   INSTRUCTIONS

1.   Produce all responsive documents prepared, sent, or received, in whole or in part, during the time period applicable to the request. You should identify in some reasonable manner the person from whose files each document you produce was found.

2.   You are to produce all documents, as defined above, that are in the possession, custody or control of you, or in the possession, custody or control of any attorney or agent for you. Without limiting the terms, "possession, custody, or control" of an item means that you either have physical possession of the item or have a right to possession of the item that is equal or superior to the person who has physical possession of the item.

3.   You are under a duty to supplement your responses to these discovery requests pursuant to FRCP 26. If after you produce any of the discovery requested below you discover or receive additional documents or information responsive to the requests, you should promptly produce such additional documents or information to the full extent required by the rules.

4.   Pursuant to FRCP 26, you are to provide a privilege log for any documents you have withheld on the basis of privilege(s). Request for such a log is hereby made, effective as of the response date of these requests. If you assert a claim of privilege with respect to any document

request, or portion thereof, you should produce a log, accounting for each document withheld individually, and providing the nature of the privilege being claimed; the type of document, date, author, addressees or recipients, and general subject matter of the document, and such other information as is sufficient to identify the document. If a claim of privilege applies to only a portion of a document, the document should be produced, and the portion claimed to be privileged obscured and stamped "redacted."

5.   The documents requested below should be produced in the manner that they are kept in the usual course of business. Documents should be produced with the label or labels from any file folder, binder, file drawer, file box, notebook, computer disk, or other container in which the document was found. Documents on computers or computer disks, including but not limited to electronic mail and corresponding attachments, should be produced in both native format and single-page TIFF format with optical character recognition (OCR). Even if only part of a document is responsive to a request, you should still produce the entire document including all attachments.

6.   All duplicates or copies of documents are to be provided to the extent they have handwriting, additions, or deletions of any kind different from the original document being produced.

7.   All documents which cannot be legibly copied should be produced in their original form.

8.   If any documents requested herein have been lost, discarded or destroyed, they shall be identified as completely as possible, including, without limitation, the following information: (a) the date of disposal; (b) the manner of disposal; (c) the reason for disposal; (d) the person authorizing disposal; and (e) the person disposing of the document.

9.   If no Documents responsive to a particular request exist, state that no responsive Documents exist.

10.   If you or your lawyers find any of these requests vague, confusing, hard to understand, or if they just want to talk through some issues relating to the requests, please call Elisha Barron at 212-729-2013. Please do not wait and object instead of attempting to resolve the issue with a telephone call. Thank you.

III.   DOCUMENT REQUESTS

REQUEST FOR PRODUCTION NO. 1:

Documents concerning the use of Times Content for training or fine tuning Your Text Generation AI Models, including, for every Text Generation AI Model that uses Times Content: i) the amount and prevalence of Times Content in the Training Datasets by token count, word count, article count, as a percentage of the Training Datasets, or by any other metric; ii) rules or guidance for choosing what is and is not included in the Training Datasets for the Text Generation AI Models; iii) use of Times Content to clean or curate the Training Datasets; iv) any weighting of particular Training Datasets for the Text Generation AI Models; and v) use of Times Content for any post-training activity, including fine tuning.

REQUEST FOR PRODUCTION NO. 2:

Documents You have gathered to submit or have submitted to any legislative or executive agency, committee, or other governmental entity in the United States that concern or relate to the allegations in the Complaint, including any Documents related to Your intended or actual submission.

REQUEST FOR PRODUCTION NO. 3:

Documents concerning copyrighted material in the Text Generation AI Models, the Training Datasets, or any Generative AI Products and Services, including documents or communications regarding the appropriateness, value, importance, necessity, and/or concerns over or legal implications of the use of such materials.

REQUEST FOR PRODUCTION NO. 4:

Documents concerning claims of copyright and trademark infringement, including complaints regarding infringement, steps taken to respond to such claims, and steps taken to prevent infringement.

REQUEST FOR PRODUCTION NO. 5:

Documents reflecting policies, procedures or practices concerning use of intellectual property in AI Models, including the policy OpenAI referred to as "Copyright Shield" in the November 6, 2023, blog post on Your website titled "New models and developer products announced at DevDay."

REQUEST FOR PRODUCTION NO. 6:

Documents concerning alternatives to using copyrighted content to train the Text Generation AI Models without compensation including using licensing agreements or material in the public domain.

REQUEST FOR PRODUCTION NO. 7:

Documents concerning the impact of AI Models on Journalism and the past, current or future use of the Text Generation AI Models or any Generative AI Products and Services as a source of Journalism, substitute for Journalism, or product in competition with Journalism (including The Times's Journalism).

REQUEST FOR PRODUCTION NO. 8:

Documents sufficient to show to what extent and in what form Times Content resides in, is stored in, is copied by, or otherwise exists within the Training Datasets and/or the Text Generation AI Models.

REQUEST FOR PRODUCTION NO. 9:

Documents sufficient to show and describe OpenAI's corporate organizational structure,

including those sufficient to show OpenAI's board of directors, parent companies, subsidiaries, affiliates, departments, and divisions; and those sufficient to show the function each Defendant plays within Your corporate structure as a whole.

REQUEST FOR PRODUCTION NO. 10:

Documents sufficient to identify OpenAI's current and former officers, directors, managers, employees and consultants with knowledge of the Text Generation AI Models, business practices regarding the Text Generation AI Models, and commercialization and use of the Text Generation AI Models.

REQUEST FOR PRODUCTION NO. 11:

Documents concerning OpenAI's transition to a for-profit company, as alleged in paragraphs 57-58 of the Complaint.

REQUEST FOR PRODUCTION NO. 12:

Documents concerning Defendants' use of Retrieval Augmented Generation ("RAG") in Generative AI Products and Services, including: (i) use of Times Content for RAG; (ii) rules or guidance for preferring or avoiding certain web domains (including The Times's web domains or sites known to host content copied from other sources) when choosing what to include in a generative result; (iii) rules for whether and how the results from RAG search products are similar to or different from standard search results; (iv) and rules regarding paywalls.

REQUEST FOR PRODUCTION NO. 13:

Documents concerning the features or capabilities of Retrieval Augmented Generation.

REQUEST FOR PRODUCTION NO. 14:

A complete listing of every data sample in the Training Datasets used to train each of the Text Generation AI Models and the source from which each data sample was obtained.

REQUEST FOR PRODUCTION NO. 15:

Documents concerning any commercial or supercomputing collaborations between and among the Defendants relating to the Generative AI Models, including licensing, developing, training, hosting, engineering, testing, deploying, or operating Generative AI Product(s) and Service(s).

Dated: February 23, 2024   /s/ Elisha Barron

Ian Crosby *(admitted pro hac vice)*
Genevieve Vose Wallace *(admitted pro hac vice)*
Katherine M. Peaslee *(pro hac vice pending)*
SUSMAN GODFREY L.L.P.
401 Union Street, Suite 3000
Seattle, WA 98101
Telephone: (206) 516-3880
Facsimile: (206) 516-3883
icrosby@susmangodfrey.com
gwallace@susmangodfrey.com
kpeaslee@susmangodfrey.com

Davida Brook *(admitted pro hac vice)*
Emily K. Cronin (*admitted pro hac vice*)
Ellie Dupler *(admitted pro hac vice)*
SUSMAN GODFREY L.L.P.
1900 Ave of the Stars, Suite 1400
Los Angeles, CA 90067
Telephone: (310) 789-3100
Facsimile: (310) 789-3150
dbrook@susmangodfrey.com
ecronin@susmangodfrey.com
edupler@susmangodfrey.com

Elisha Barron (5036850)
Zachary B. Savage (ZS2668)
Tamar Lusztig (5125174)
Alexander Frawley (5564539)
Eudokia Spanos (5021381)
SUSMAN GODFREY L.L.P.
1301 Avenue of the Americas, 32nd Floor
New York, NY 10019
Telephone: (212) 336-8330
Facsimile: (212) 336-8340

ebarron@susmangodfrey.com
zsavage@susmangodfrey.com
tlusztig@susmangodfrey.com
afrawley@susmangodrey.com
espanos@susmangodfrey.com

Scarlett Collings *(Admission Pending)*
SUSMAN GODFREY L.L.P.
1000 Louisiana, Suite 5100
Houston, TX 77002
Telephone: (713) 651-9366
Facsimile (713) 654-6666
scollings@susmangodfrey.com

Steven Lieberman (SL8687)
Jennifer B. Maisel (5096995)
Kristen J. Logan *(admitted pro hac vice)*
ROTHWELL, FIGG, ERNST & MANBECK, P.C.
901 New York Avenue, N.W., Suite 900 East
Washington, DC 20001
Telephone: (202 783-6040
Facsimile: (202) 783 6031
slieberman@rothwellfigg.com
jmaisel@rothwellfigg.com
klogan@rothwellfigg.com

Attorneys for Plaintiff
The New York Times Company